AD-757 168

MAXIMUM LIKELIHOOD ESTIMATION FROM RENEWAL TESTING

Larry H. Crow

Army Materiel Systems Analysis Agency
Aberdeen Proving Ground, Maryland

April 1972

AD752168

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) U.S. Army Materiel Systems Analysis Agency Aberdeen Proving Ground, Maryland	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

MAXIMUM LIKELIHOOD ESTIMATION FROM RENEWAL TESTING

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

Larry. H. Crow

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1972	68	12

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 57
b. PROJECT NO. AMCMS Code 2270.1031	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	U.S. Army Materiel Command Washington, D.C.

13. ABSTRACT

This report considers the maximum likelihood estimates of life-time distributions over an interval [0,T) from the following time truncated experiment. At time zero, the beginning of the testing $n(1 \leq n < \infty)$ items are put on test. When an item fails it is replaced and at time T all testing is stopped.

Assumptions about the form of the life-time distribution on [0,T) are required. Distributions considered are:

(1) A single parameter class which includes the Weibull family;

(2) A multiple parameter class with increasing failure rate on [0,T); and

(3) A nonparametric class which includes the increasing failure rate family.

Useful and desirable properties of the maximum likelihood estimates are shown.

1-a

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

Unclassified

Security Classification

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Maximum likelihood estimation Renewal life-testing Weibull family Increasing failure rate Asymptotic normality Strong consistency						

1-4

Unclassified
Security Classification

TECHNICAL REPORT NO. 57

MAXIMUM LIKELIHOOD ESTIMATION FROM RENEWAL TESTING

Larry H. Crow

April 1972

AMCMS Code 2270.1031

U.S. ARMY MATERIEL SYSTEMS ANALYSIS AGENCY
ABERDEEN PROVING GROUND, MARYLAND

2

U.S. ARMY MATERIEL SYSTEMS ANALYSIS AGENCY

TECHNICAL REPORT NO. 57

LHCrow/da
Aberdeen Proving Ground, Md.
April 1972

MAXIMUM LIKELIHOOD ESTIMATION FROM RENEWAL TESTING

ABSTRACT

This report considers the maximum likelihood estimates of life-time distributions over an interval $[0,T)$ from the following time truncated experiment. At time zero, the beginning of the testing $n(1 \leq n < \infty)$ items are put on test. When an item fails it is replaced and at time T all testing is stopped.

Assumptions about the form of the life-time distribution on $[0,T)$ are required. Distributions considered are:

(1) A single parameter class which includes the Weibull family;

(2) A multiple parameter class with increasing failure rate on $[0,T)$; and

(3) A nonparametric class which includes the increasing failure rate family.

Useful and desirable properties of the maximum likelihood estimates are shown.

Next page is blank.

CONTENTS

Next page is blank.

MAXIMUM LIKELIHOOD ESTIMATION FROM RENEWAL TESTING

1. INTRODUCTION AND SUMMARY

Based on a number of practical reasons it is often necessary and even desirable in life testing (reliability) studies to fix the total testing time, say, at T ($T < \infty$), before testing begins. For example, an experimenter would rarely use a testing plan that did not limit the total testing time when the items being tested can be assumed very reliable, since the testing time would usually be very long. The total testing time must, also, be limited if project deadlines must be met, or if equipment or personnel used in the testing can only be spared for some specified length of time.

Limiting the total testing time need not, however, be contrary to the goals of the experimenter. For example, if the experimenter can assume that the general form of the life-time distribution belongs to some parametric class defined on the nonnegative real axis, then limiting the testing time to T, he can still estimate the unknown parameters of the distribution on $[0, \infty)$.

If the experimenter cannot assume that the life-time distribution has a particular form on $[0, \infty)$ but only on $[0, T)$, then he must limit his inferences to the latter interval. However, if $[0, T)$ includes the mission time of the items tested then, for all practical purposes, he need not infer anything about the distribution outside this interval.

One of the most popular time truncated testing plans is the subject of this report. This plan stipulates that n items are initially put on test at time zero. When an item fails it is replaced by a new item and at time T all testing is stopped. (For convenience this plan is called "Testing Plan A.") Renewing a failed item is a method to further save experimental time and generally results in a better utilization of equipment and personnel.

Practically all of the statistical procedures developed in the literature for Testing Plan A are based on the assumption that the underlying life-time distribution of the items tested is the exponential law

$$G(x) = 1 - \exp(-\lambda x), \tag{1.1}$$

$\lambda > 0$, $x > 0$. (See Epstein [1959] for a review of these procedures.) In practice, however, the exponential assumption is often not valid since it implies a no wear-out (or no aging) property of the items. Moreover, if the times to failure of the items do not follow the law (1.1) these exponential procedures could possibly be sensitive to this departure. (See Zelen and Dannemiller [1961].)

For Testing Plan A this report investigates the maximum likelihood estimates of life-time distributions from three general classes. Life-time distributions describing wear-out are contained in each of these classes and, also, each class contains the exponential distribution.

Specifically, in Section 3 the maximum likelihood estimate (MLE) of the parameter λ will be considered when the life-time distribution has the form

$$F(x) = 1 - \exp(-\lambda g(x)),$$

$\lambda > 0$, $0 \leq x < T$, $g(\cdot)$ is a known, strictly increasing, differentiable function on $[0, T)$ with $g(0) = 0$. Observe that nothing is assumed about F on $[T, \infty)$. This parametric class is obviously relevant to life testing since, for example, it includes the exponential (when $g(x) = x$, $x > 0$), the Weibull (when $g(x) = x^{\beta}$, $\beta > 0$, $x > 0$) and the extreme-value distributions (when $g(x) = e^{x} - 1$, $x > 0$). Asymptotic distribution theory, which will allow one to test hypothesis on the true value of λ, shall be given along with a number of pleasant properties of the MLE. These results do not depend on the fact that F is not restricted on $[T, \infty)$. Also, a major drawback to another method of estimating λ shall be discussed.

As such, the class of distributions introduced in Section 4 has not been considered in the literature. Practical applications of this class shall be discussed and the MLE's of parameters determining the life-time distributions over $[0, T)$ are shown to be asymptocially normal and consistent.

Often an experimenter does not know a priori that the law governing the times to failure of the items tested belongs to a certain parametric class. He may, however, know that the underlying distribution is a member of a non-parametric class of distribution, e.g., the Increasing (Decreasing) Failure Rate (IFR (DFR)) family.

Marshall and Proschan (1965) considered the MLE of a life-time distribution, assuming only that it was a member of the IFR (DFR) family and that data arise from a testing plan which does not allow censoring, time-truncation or replacement. Bray, Crawford and Proschan (1967), also, considered the MLE of a life-time distribution from a nonparametric class which includes both the IFR and DFR families. The testing plan they introduced allowed for the consideration of various types of incomplete data.

Since nonparametric estimation has not been considered in the literature for Testing Plan A, we will study this type of estimation in Section 5. The class of distributions considered includes the IFR family and the main result of that section is the consistency of the MLE over $[0, T)$.

2. PRELIMINARIES

In this section preliminary definitions and notations needed in later sections shall be collected. For completeness we give

Definition 2.1 (Testing Plan A)

At time zero, the beginning of the testing, n new items from a population are put on test. When an item fails it is instantaneously replaced with a new item from the original population and at time T the testing is stopped.

Estimation from Testing Plan A has been considered by several authors, including Epstein (1959, page 3.17) and Gnedenko, Belyayev and Solovyev (1969, page 169), when the life-time distribution of the items is exponential. These authors derived the likelihood function by standard methods which could, also, be used for other classes of distributions with densities. In the present approach, however, the derivation of the likelihood function utilizes the theory of stopping variables. The benefits of this approach are two-fold. Firstly, a straightforward method of obtaining the likelihood function is developed for the parametric classes of distributions considered in Sections 3 and 4. Finally, this approach motivates a generalized likelihood function needed in Section 5 for the nonparametric class.

To develop this preliminary theory observe that Testing Plan A may be considered as n independent experiments, each beginning at time zero and ending at time T. Throughout this report K_r will denote the random number of items put on test in the r-th experiment and X_{ir} will denote the time to failure of the i-th item put on test in this experiment, $i=1,2,\ldots,r=1,\ldots,n$. From this notation we have that K_r is the first integer such that

$$\sum_{i=1}^{K_r} X_{ir} \geq T,$$

$r=1,\ldots,n$.

For the moment consider only the 1-st experiment and let $X_i = X_{i1}$, $i=1,2,\ldots,K=K_1$. Also, let F be the cumulative distribution

function (c.d.f.) of X_1. It is straightforward to show that if $F(0) < 1$ then K is exponentially bounded; i.e.,

$\text{Prob}(K=k) \leq cp^k, c > 0, 0 < p < 1, k = 1,2,\ldots$.

This implies that K is finite with probability one (w.pr.1) and all moments of K exist. Now, let

$$A_k = \{x_1,\ldots,x_{k-1}: \sum_{i=1}^{k-1} x_i < T\}$$

and

$$B_k = \{x_1,\ldots,x_k: \sum_{i=1}^{k-1} x_i < T, \sum_{i=1}^{k} x_i \geq T\}$$

$k=1,2,\ldots$. (The convention that sums of the form $\sum_{i=1}^{0}$ are equal to 0 and products of the form $\prod_{i=1}^{0}$ are equal to 1 is adopted in this report.) Since $K < \infty$ w.pr.1, one may show that

$$1 = \sum_{k=1}^{\infty} \int \cdots \int_{B_k} \prod_{i=1}^{k} dF(x_i). \tag{2.1}$$

Let Y_{ir} be the time on test of the i-th item put on test in the r-th experiment, $i=1,2,\ldots,r=1,\ldots,n$, and let $Y_i = Y_{i1}, i=1,2,\ldots$. Observe, now, that

$$Y_{ir} = X_{ir}, \tag{2.2}$$

$i=1,\ldots,K_r-1$, and

$$Y_{K_r,r} = T - \sum_{i=1}^{K_r-1} X_{ir}, \tag{2.3}$$

$r=1,\ldots,n$. Since the testing is truncated at time T, experiment 1 is characterized by the time on test statistics $(Y_1,\ldots,Y_K)$. By (2.2) and (2.3) one sees that the experiment is equivalently characterized by the times to failure $(X_1,\ldots,X_{K-1})$.

Now, in almost all cases where the c.d.f. has a probability density function (p.d.f) one can show that the likelihood function is derived from the integrand of an expression equated to 1 and where the integration is over the sample space of the random variables of interest. For $(X_1,\ldots,X_K)$ the sample space is

$$\Omega_1 = \bigcup_{k=1}^{\infty} B_k,$$

and for $(X_1,\ldots,X_{K-1})$ the sample space is

$$\Omega_2 = \bigcup_{k=1}^{\infty} A_k.$$

We, therefore, integrate out x_k in (2.1) obtaining

$$1 = \sum_{k=1}^{\infty} \int_{A_k} \cdots \int [1-F(\{T - \sum_{i=1}^{k-1} x_i\}^-)] \prod_{j=1}^{k-1} dF(x_i), \tag{2.4}$$

where

$$F(x^-) = \lim_{\varepsilon \to 0} F(x-\varepsilon),$$

$\varepsilon>0, x>0$, and it is observed that

$$B_k = \{x_1,\ldots,x_k : \sum_{i=1}^{K_r-1} x_i < T,\ x_k \geq T - \sum_{i=1}^{k-1} x_i\}$$

$k=1,2,\ldots$.

Hence, if F is absolutely continuous on $[0, T)$ with p.d.f. f, then

$$1 = \sum_{k=1}^{\infty} \int \ldots \int_{A_k} [1-F(\{T - \sum_{i=1}^{k-1} x_i\}^-)] \prod_{j=1}^{k-1} f(x_j)\, dx_j. \tag{2.5}$$

This motivates the following.

Definition 2.2.

If the times to failure of the items are independent and identically distributed (iid) with c.d.f. F, $F(0) = 0$, and F is absolutely continuous on $[0, T)$ with p.d.f. f, then the likelihood function L for Testing Plan A is

$$L = \prod_{r=1}^{n} L_r, \tag{2.6}$$

where

$$L_r \equiv L_r(X_{1r}, \ldots, X_{K_r-1,r})$$

is the random variable,

$$L_r = [1-F(\{T - \sum_{i=1}^{K_r-1} X_{ir}\}^-)] \prod_{j=1}^{K_r-1} f(X_{jr}). \tag{2.7}$$

Equation (2.6) is a result of the independence of the n experiments and Equation (2.7) is obtained from the integrand of Equation (2.5) when the random variables replace their corresponding sample points. Definition 2.2 will be used in Section 3 and 4 to derive the MLE's for the parametric classes of distributions. A generalized definition of MLE, based on Equation (2.4), will be defined in Section 5 for the nonparametric class.

3. SINGLE PARAMETER ESTIMATION

3.1 Introduction.

Throughout this section it will be assumed that the underlying c.d.f. of the times to failure is

$$F(x) = 1 - \exp(-\lambda g(x)) \tag{3.1}$$

for $0 < x < T$, $\lambda > 0$, $g(\cdot)$ is known and strictly increasing with $g(0) = 0$ and derivative $g'(x)$ $0 < x < T$. We shall derive the MLE $\hat{\lambda}_n$ of λ and show that: (a) $\hat{\lambda}_n$ is strongly consistent, as $n \to \infty$; (b) $\hat{\lambda}_n$ is asymptotically normally distributed as $n \to \infty$; (c) $\hat{\lambda}_n$ is asymptotically efficient, as $n \to \infty$. Also, a major drawback to some previously published work dealing with the estimation of λ shall be discussed.

3.2 Previous Work.

Gnedenko, Belyayev and Solovyev (1969), devoted an entire section of their book, beginning on page 168, to the MLE of the parameter λ when the underlying distribution is the exponential law

$$G(x) = 1 - \exp(-\lambda x) \tag{3.2}$$

$x > 0$, $\lambda > 0$, for six life testing plans, one of which was Testing Plan A. Observe, now, that if X is a random variable with c.d.f. given by (3.1) for $x > 0$ then $g(X)$ is a random variable with c.d.f. given by (3.2). Noting this, Gnedenko, et al mentioned that if the life-time distribution is given by (3.1) for $x > 0$ then one may make the transformation $Y = g(X)$ on the data and use their exponential procedures to estimate λ. However, it was not pointed out that if the exponential procedures are used for Testing Plan A then the total testing time will not necessarily be T, which violates the purpose of this testing plan.

To see this difficulty, observe that the suggested test plan implies that one make the transformation

$$W_{ir} = g(X_{ir})$$

$i \geq 1$, r=1, ..., n, choose a constant $C > 0$, and continue testing in the r-th experiment (r=1, ..., n) until time C on the $g(\cdot)$ time axis. If one does this then the random number of items, K_r, put on test in the r-th experiment is the first integer such that

$$\sum_{i=1}^{K_r} W_{ir} \geq C,$$

r=1, ..., n. Therefore, the actual (untransformed) total testing time in the r-th experiment is

$$\sum_{i=1}^{K_r-1} X_{ir} + g^{-1}\left(C - \sum_{i=1}^{K_r-1} g(X_{ir})\right),$$

r=1, ..., n. If C is to be chosen such that the total testing time is T, then for $K_r = 1$, the total testing time is $g^{-1}(C) = T$. Thus, $g(T)$ is the only candidate for C. Now, if

$$\sum_{i=1}^{K_r-1} X_{ir} + g^{-1}\left(g(T) - \sum_{i=1}^{K_r-1} g(X_{ir}) \right) = T$$

for $K_r > 1$, this would imply that

$$g(T) - \sum_{i=1}^{K_r-1} g(X_{ir}) = g(T - \sum_{i=1}^{K_r-1} X_{ir}),$$

which is, in general, not true for non-linear g. Thus, when one makes such a transformation the total testing times for the n experiments will generally be random variables. This violates the purpose of Testing Plan A which is to fix the total testing time at T. The work presented in this section allows one to estimate λ without using such a transformation and, hence, avoiding this difficulty.

3.3 The MLE of λ.

The MLE $\hat{\lambda}_n$, say, of λ shall now be derived. In what follows let

$$g(T) = \lim_{\varepsilon \to 0} g(T-\varepsilon), \ \varepsilon > 0.$$

Lemma 3.1

The MLE $\hat{\lambda}_n$ of λ is

$$\lambda_n = \sum_{r=1}^{n} (K_r - 1) / \sum_{r=1}^{n} \sum_{i=1}^{K_r} g(Y_{ir}). \tag{3.3}$$

Proof:

By Definition 2.2 the likelihood function is

$$L = \prod_{r=1}^{n} L_r$$

where

$$L_r = \exp(-\lambda g(Y_{K_r r})) \prod_{i=1}^{K_r - 1} \lambda g'(Y_{ir}) \exp(-\lambda g(Y_{ir})),$$

$r=1,\ldots,n$. Maximizing L with respect to λ yields $\hat{\lambda}_n$ given by (3.3).

The reader should note that if F is continuous at T, then

$$\sum_{r=1}^{n} (K_r - 1)$$

is the number of failures in the n experiments. Also, if the times to failure are exponentially distributed (i.e., $g(x) = x$, $x > 0$), then

$$\sum_{r=1}^{n} \sum_{i=1}^{K_r} g(Y_{ir}) = nT.$$

Hence, $\hat{\lambda}_n$ is the usual estimator in the exponential case.

Observe, now, that identity (2.4) implies that the probability of <u>any event</u> associated with the outcome of an experiment based on Testing Plan A only depends on $F(x)$ for $0 < x < T$. Since these probabilities do not depend on $F(x)$ for $T \leq x < \infty$, it follows that the statistical properties of any random variable obtained from Testing Plan A are independent of $F(x)$ for $T \leq x < \infty$. We, therefore, have

<u>Theorem 3.2</u>

The statistical properties of λ_n and all other random variables obtained from Testing Plan A do not depend on the values of F on $[T, \infty)$.

3.4 <u>Strong Consistency of $\hat{\lambda}_n$.</u>

We will now show that $\hat{\lambda}_n$ converges to λ almost surely*(a.s.) as $n \to \infty$. To show this we will need the following results.

<u>Lemma 3.3</u>

If F_1 is any c.d.f. such that $F(x) = F_1(x)$, $0 \leq x < T$, then

$$P[K_1 = k|F] = P[K_1 = k|F_1], \text{ for all } k = 1, 2, \ldots$$

<u>Proof:</u>

The proof follows from Theorem 3.2.

*The term "almost surely" means that a certain event holds with probability one.

The following result is needed to show consistency and is, also, useful throughout the remainder of this section.

Theorem 3.4

$$\frac{1}{\lambda} E(K_1-1) = E(\sum_{i=1}^{K_1} g(Y_{i1})). \tag{3.4}$$

Proof:

Let $X_i = X_{i1}$, $i = 1,2,\ldots, K = K_1$. By Lemma 3.3 $E(K)$ does not depend on $F(x)$ for $x \geq T$. Hence, if $g(T) < \infty$ then we may extend $g(x)$ for $x \geq T$ in any manner we wish to keep F a c.d.f. and $E(K)$ will remain unchanged. We therefore assume that $g(x) = g(T) + (x-T)$ for $x \geq T$, when $g(T) < \infty$. Hence, whether or not $F(T) < 1$ or $F(T) = 1$, $g(X_i)$ has an exponential distribution with mean $1/\lambda$. By Wald's Lemma (1944)

$$\frac{1}{\lambda} E(K) = E(\sum_{i=1}^{K} g(X_i)). \tag{3.5}$$

Now,
$$\begin{aligned} E(g(X_K)) &= E(E(g(X_K)|\sum_{i=1}^{K-1} X_i)) \\ &= E(E(g(X) \;|X \geq T - \sum_{i=1}^{K-1} X_i)) \\ &= E(E(g(X)|g(X) \geq g(T - \sum_{i=1}^{K-1} X_i))) \end{aligned}$$

where $g(X)$ is a random variable with c.d.f. $1-\exp(-\lambda y)$, $y > 0$. Thus,

$$\begin{aligned} E(g(X_K)) &= E(\frac{1}{\lambda} + g(T - \sum_{i=1}^{K-1} X_i)) \\ &= \frac{1}{\lambda} + E(g(T - \sum_{i=1}^{K-1} X_i)). \end{aligned} \tag{3.6}$$

Equations (3.5) and (3.6) imply (3.4).

We may now prove

Theorem 3.5

The MLE, $\hat{\lambda}_n$, given by Equation (3.3), is a strongly consistent estimator of λ as $n \to \infty$.

Proof:

By the strong law of large numbers, as $n \to \infty$

$$\sum_{r=1}^{n} (K_r-1)/n \longrightarrow E(K_1-1) \text{ a.s.},$$

and

$$\sum_{r=1}^{n} \sum_{i=1}^{K_r} g(Y_{ir})/n \longrightarrow E(\sum_{i=1}^{K_1} g(Y_{i1})) \text{ a.s. .}$$

Hence, as $n \to \infty$

$$\hat{\lambda}_n = \frac{\sum_{r=1}^{n} \sum_{i=1}^{K_r} (K_r-1)/n}{\sum_{r=1}^{n} \sum_{i=1}^{K_r} g(Y_{ir})/n} \longrightarrow \frac{E(K_1-1)}{E(\sum_{i=1}^{K_1} g(Y_{i1}))} \text{ a.s. .}$$

The result follows from Equation (3.4).

3.3 Asymptotic Normality of $\hat{\lambda}_n$.

We will now show the asymptotically normality of the MLE $\hat{\lambda}_n$ for two different, but asymptotically equivalent, normalizing sequences. We begin with

Theorem 3.6

The asymptotic distribution of $(\hat{\lambda}_n-\lambda)/\sqrt{D/n}$ is Normal (0,1), as $n \to \infty$ where

$$D = \frac{\operatorname{Var}((K_1-1) - \lambda \sum_{j=1}^{K_1} g(Y_{j1}))}{E^2(\sum_{j=1}^{K_1} g(Y_{j1}))}. \qquad (3.7)$$

Proof:

$$\text{Let } M_r = \sum_{j=1}^{K_r} g(Y_{jr}),\ V_r = (K_r-1),\ M(n) = \sum_{r=1}^{n} \frac{M_r}{n},\ V(n) = \sum_{r=1}^{n} \frac{V_r}{n},$$

and let $\underset{\sim}{Z}_r$ be the two dimensional random vector, $\underset{\sim}{Z}_r = (M_r, V_r)$, $r = 1,\ldots,n$. Also, let $H(a,b)$ be the function of the two variables a, b, $H(a,b) = a/b$. Now, $(M(n), V(n))$ is the first moment vector corresponding to the sample $\underset{\sim}{Z}_1, \underset{\sim}{Z}_2, \ldots, \underset{\sim}{Z}_n$. By Cramér (1946, pages 353, 367), $H(V(n), M(n)) = \hat{\lambda}_n$ is asymptotically normal with asymptotic mean $\frac{E(V(n))}{E(M(n))} = \lambda$ by Equation (3.4), and asymptotic variance

$$\frac{\text{Var}(V(n))}{E^2(M(n))} - 2\,\text{Cov}(V(n), M(n))\,\frac{E(V(n))}{E^3(M(n))} + \text{Var}(M(n))\,\frac{E^2(V(n))}{E^4(M(n))}\ .$$

Using Equation (3.4) again the asymptotic variance equals

$$\frac{1}{nE^2\left(\sum_{j=1}^{K_1} g(Y_{j1})\right)}\left[\text{Var}(K_1-1) - 2\lambda\,\text{Cov}\left(K_1-1, \sum_{j=1}^{K_1} g(Y_{j1})\right) + \lambda^2\,\text{Var}\left(\sum_{j=1}^{K_1} g(Y_{j1})\right)\right].$$

This completes the proof.

The next theorem will be useful in what follows.

Theorem 3.7

$$E(K_1-1) = \text{Var}\left(K_1-1-\lambda \sum_{j=1}^{K_1} g(Y_{j1})\right) \tag{3.8}$$

Proof:

Let $K = K_1$, $X_i = X_{i1}$, $i = 1,\ldots, K-1$, and $Y_i = Y_{i1}$, $i = 1,\ldots,K$. Also, let $f(x) = \lambda g'(x)\exp(-\lambda g(x))$, and

$$p(x_1,\ldots,x_{k-1}|\lambda) = [1-F(\{T-\sum_{i=1}^{k-1} x_i\}^-]\prod_{j=1}^{k-1} f(x_j).$$

It is easy to verify that

$$E\left(\frac{d}{d\lambda}\log p(X_1,\ldots,X_{K-1}|\lambda)\right)^2 = -E\left(\frac{d^2}{d\lambda^2}\log p(X_1,\ldots,X_{K-1}|\lambda)\right) . \tag{3.9}$$

The left-hand side of (3.9) is equal to

$$E\left(\frac{K-1}{\lambda} - \sum_{j=1}^{K} g(Y_j)\right)^2.$$

But using Equation (3.4) we have

$$\mathrm{Var}\left(\frac{K-1}{\lambda} - \sum_{j=1}^{K} g(Y_j)\right) = E\left(\frac{K-1}{\lambda} - \sum_{j=1}^{K} g(Y_j)\right)^2. \tag{3.10}$$

The right-hand side of (3.9) is equal to $E(K-1)/\lambda^2$. Hence, (3.8) follows.

Using Equations (3.4) and (3.8) it follows, also, that

$$D = \lambda^2/E(K_1-1). \tag{3.11}$$

From this we have

<u>Corollary 3.8</u>

The asymptotic distribution of $(\hat{\lambda}_n-\lambda)/\sqrt{\lambda^2/(nE(K_1-1))}$ is Normal (0,1) as $n \to \infty$.

By the strong law of large numbers Corollary 3.8 gives

<u>Corollary 3.9</u>

The asymptotic distribution of $(\hat{\lambda}_n-\lambda)/\sqrt{\lambda^2/(\sum_{r=1}^{n} (K_r-1)}$ is Normal (0,1) as $n \to \infty$.

3.6 Asymptotic Efficiency of $\hat{\lambda}_n$.

Let $h(X_1,\ldots,X_{K-1})$ be an estimate of λ, where $X_i = X_{i1}$, $i = 1,\ldots, K-1$, $K = K_1$. Then

<u>Theorem 3.10</u>

$$D\left(1 + \frac{d}{d\lambda} B(\lambda)\right)^2 \leq \mathrm{Var}(h(X_1,\ldots,X_{K-1})),$$

where $B(\lambda) = E(h(X_1,\ldots,X_{K-1})|\lambda) - \lambda$, and D is given by Equation (3.7).

Proof:

It is straightforward to show that

$$[1 + \frac{d}{d\lambda} B(\lambda)]^2 \leq \mathrm{Var}(h(X_1,\ldots,X_{K-1})|\lambda)\ E(\frac{d}{d\lambda} \log p(X_1,\ldots,X_{K-1}|\lambda))^2.$$

(Note that in the sequential form of the Cramér-Rao bound that $\prod_{i=1}^{k} f(x_i)$ corresponds to $p(x,\ldots,x_{k-1}|\lambda)$.) Using Equations (3.8), (3.10), and (3.11) yields the result.

This implies that if h_1 is an unbiased estimator of λ based on the outcome of n experiments, then

$$\frac{D}{n} \leq \mathrm{Var}(h_1). \tag{3.12}$$

From this we have the following

Theorem 3.11

$\hat{\lambda}_n$ is an asymptotically efficient estimator of λ.

Proof:

Our concept of efficiency is the same as the concept given by BAN estimators for fixed sample size. (See Rao (1968), p. 284.) The result then follows from Theorem 3.6 and inequality (3.12).

3.7 Comments.

Note that if the times to failure of the items put on test actually have the c.d.f. $F(x) = 1 - \exp(-\lambda g(x))$ for $0 < x < T + b$, $0 < b \leq \infty$, then, of course, the assumptions required for $F(\cdot)$ are satisfied. In this case

the estimate $\hat{\lambda}_n$ allows one to estimate $F(x)$ for $0 < x < T + b$ from data restricted to $[0,T]$. Suppose, however, that the c.d.f. has the form $F(x) = 1 - \exp(-\lambda g(x))$, $0 < a \leq x < T + a$. Then items of age a have the c.d.f. $G(x) = 1 - \exp(-\lambda q(x))$ $a \leq x < T + a$, where $q(x) = g(x) - g(a)$. Thus, one may put items of age a on test at time 0 and use the theory presented in this section to estimate $F(x)$ for $a \leq x < T + a$.

It is to be remarked, also, that numerous computer simulation runs substantiate the conjecture that the MLE of λ is generally not unbiased; i.e., in general $E(\hat{\lambda}_n) \neq \lambda$. However, the bias approaches zero as n or T gets large.

4. MULTIPLE PARAMETER ESTIMATION

4.1 Introduction.

In this section it is assumed that the c.d.f. F governing the times to failure of the items put on test is absolutely continuous on $[0,T)$ with p.d.f. f and $F(0) = 0$. Also, it is assumed that the failure rate $f(x)/[1-F(x)] = \lambda_q$, for $x \in [S_q, S_{q+1})$, $q = 0,1,\ldots, t-1$, where $0 = S_0 < S_1 < \ldots < S_t = T$, and $0 \leq \lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_{t-1} < \infty$. Thus, assuming that the S_q, $q = 0,1,\ldots,t$ are known, and data are collected from Testing Plan A, the MLE's of λ_q, $q = 0,1,\ldots, t-1$, are determined and shown to be strongly consistent estimators as $n \to \infty$. The asymptotic normality of these estimators, as $n \to \infty$ is, also, established.

Consider now a situation when Testing Plan A and this class of distributions may be applicable. The guidance system or some other system or component in a rocket may have a failure rate which is constant when the booster of the first stage of the rocket is in operation. However, when the first stage falls away and the second stage booster is fired the failure rate of the system may change and in fact increase instantly to a constant value during this stage. If this is true for all stages of the rocket, then, (since the exact length of each stage and the exact time of the staging is known), laboratory

testing may be used to estimate the failure rate of the system for the duration of its mission.

Another possible application may arise when one is interested in estimating the failure rate of an electronic apparatus as a function of the amount of voltage. It is not unusual for electronic tubes and the like to have a constant failure rate when the voltage is constant. If the failure rate is a nondecreasing function of the voltage then one may estimate the failure rate for specific values of the voltage in the following way. Let the testing time T be fixed and let $v_0 < v_1 < \ldots < v_{t-1}$ be voltages which are of interest to the experimenter. Let λ_i, $i = 0, \ldots, t-1$ be the failure rate of the items when they are receiving voltage v_i. Also, let $[S_i, S_{i+1})$, $i = 0, \ldots, t-1$ be a partition of $[0,T)$. When an item is put on test it receives voltage v_0. If it operates without failure for time S_1 then the voltage is increased instantly to v_1. Similarly, if the item operates for time S_i, $i < t$, then the voltage is increased to v_i. When an item fails it is replaced instantly by another new item and the voltage is reduced to v_0. If this item operates for time S_1 without failure then the voltage is increased to v_1, and so on. This process is continued until time T. The theory presented in this section will allow one to estimate the λ_i, $i = 0, \ldots, t-1$.

Applications of this model may also be possible in the fields of drug testing and toxicology. For example, suppose one

is interested in the effect of a toxic agent such as DDT or the effect of radiation, which decompose at a very slow rate. The failure rate depends on the dosage level and may be taken as constant for reasonably short periods of time and nondecreasing as the dosage level increases. The dosage is sequentially increased at the end of these successive periods and the model presented in this section may be used to estimate the failure rates corresponding to the different dosage levels for the time period of interest.

4.2 The Naive MLE of the λ_q.

We begin this section by finding the values of λ_q, $q = 0, \ldots, t-1$, which will maximize L, given by Definition 2.2, without the restriction that $\lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{t-1}$.

From Definition 2.2 the likelihood L_r for the r-th experiment is

$$L_r = [1-F((T - \sum_{i=1}^{K_r-1} X_{ir})^-)] \prod_{i=1}^{K_r-1} f(X_{ir}), \; r = 1, \ldots, n,$$

and,

the likelihood L for the n independent experiments is

$$L = \prod_{r=1}^{n} L_r.$$

Let $I(\cdot|R)$ be the indicator function of R. Furthermore, define the function G_q by

$$G_q(z_1,\ldots,z_p) = \sum_{i=1}^{p} I(z_i|[S_q,S_{q+1}))$$

and 0 otherwise, and define the function $A_q(z_1,\ldots,z_p)$ by

$$A_q(z_1,\ldots,z_p) = \sum_{i=1}^{p}\{(S_{q+1}-S_q)\ I(z_i|[S_{q+1},T])$$

$$+ (z_i-S_q)\ I(z_i|[S_q,S_{q+1}))\}.$$

Let $Y_{ir} = X_{ir}$, $r = 1,\ldots,n$, $i = 1,\ldots,K_r-1$, and $Y_{K_r,r} = T-\sum_{i=1}^{K_r-1} X_{ir}$, $r = 1,\ldots,n$. Denote by G_{rq}, $q=0,\ldots,t-1$, the function $G_q(Y_{1r},\ldots,Y_{K_r-1,r})$, which is the number of failures in the r-th experiment which lie in $[S_q, S_{q+1})$. Also denote by A_{rq}, $q = 0,\ldots, t-1$, the function $A_q(Y_{1r},\ldots,Y_{K_r,r})$, which is the total time on test for the r-th experiment over $[S_q, S_{q+1})$.

Observe, too, that if $r(x) = f(x)/(1-F(x))$, $x\varepsilon[0,T))$, then $F(x) = 1-\exp\{-\int_0^x r(y)dy\}$ and $f(x) = r(x)\exp\{-\int_0^x r(y)dy\}$ for $x\varepsilon[0,T)$.

The next lemma will allow us to easily find the values of the λ_q, $q = 0,\ldots, t-1$, say $\hat{\lambda}_{nq}$, $q = 0,\ldots, t-1$, which maximize $L = \prod_{r=1}^{n} L_r$, without the restriction that $\hat{\lambda}_{n0} \le \hat{\lambda}_{n1} \le \cdots \le \hat{\lambda}_{n(t-1)}$. We call $\hat{\lambda}_{nq}$ the "naive" MLE of λ_q. (The term "unrestricted" MLE is also used in the literature.)

Lemma 4.1

The likelihood L_r for the r-th experiment may be written as

$$L_r = \prod_{j=0}^{t-1} \lambda_j^{G_{rj}} \exp(-\lambda_j A_{rj}),\ r = 1,\ldots,n. \tag{4.1}$$

Proof:

First note that

$$r(x) = \sum_{j=0}^{t-1} \lambda_j\ I(x|[S_j,S_{j+1})),$$

for $x\varepsilon(0,T)$. Thus, L_r may be written as

$$L_r = \prod_{i=1}^{K_r-1} \left(\sum_{j=0}^{t-1} \lambda_j\ I(Y_{ir}|[S_j,S_{j+1}))\right)$$

$$\prod_{i=1}^{K_r} \exp(-\sum_{j=0}^{t-1} \lambda_j\{(S_{j+1}-S_j)\ I(Y_{ir}|[S_{j+1},T])$$

$$+ (Y_{ir}-S_j)\ I(Y_{ir}|[S_j,S_{j+1}))\}).$$

Now observe that

$$\prod_{j=0}^{t-1} \lambda_j^{G_{rj}} = \prod_{i=1}^{K_r-1} \left(\sum_{j=0}^{t-1} \lambda_j I(Y_{ir}[S_j,S_{j+1}))\right).$$

Equation (4.1) follows.

The following corollary gives the naive MLE $\hat{\lambda}_{nq}$ of λ_q, $q = 0,\ldots, t-1$.

Corollary 4.2

The maximum of $L = \prod_{r=1}^{n} L_r$ is obtained if $\lambda_q = \hat{\lambda}_{nq}$, $q = 0, \ldots, t-1$, where

$$\hat{\lambda}_{nq} = \frac{\sum_{r=1}^{n} G_{rq}}{\sum_{r=1}^{n} A_{rq}} \qquad \text{if } \sum_{r=1}^{n} A_{rq} \neq 0,$$

and

$$\hat{\lambda}_{nq} = 0 \qquad \text{if } \sum_{r=1}^{n} A_{rq} = 0.$$

Proof:

If $\sum_{r=1}^{n} A_{rq} \neq 0$, the result follows directly from (4.1).

Also, $\sum_{r=1}^{n} A_{rq} = 0$ implies that $\sum_{r=1}^{n} G_{rq} = 0$. Since $\lambda_q \geq 0$ we define 0^0 to equal 1 and thus take $\hat{\lambda}_{nq} = 0$ if $\sum_{r=1}^{n} A_{rq} = 0$. This will maximize L.

4.3 Strong Consistency of $\hat{\lambda}_{nq}$.

The next theorem will be used to show strong consistency of $\hat{\lambda}_{nq}$, $q = 0, \ldots, t-1$, given in Corollary 4.2.

Theorem 4.3

Let H be a c.d.f. sucn that $H(0) = 0$, and for u and v where $0 < u < v$, $H(v^-) - H(u^-) > 0$. Let $X_1, X_2, \ldots,$ be i.i.d. random variables with c.d.f. H, and let K be the stopping variable defined as the first integer such that

$$\sum_{i=1}^{K} X_i \geq T.$$

In addition suppose that H is absolutely continuous on $[u, v)$ with p.d.f. h and

$$h(x)/[1-H(x)] = \lambda,$$

$x \varepsilon [u, v)$. Then

$$\frac{E\left(\sum_{i=1}^{K-1} I(X_i | [u, v))\right)}{E\left(\sum_{i=1}^{K} \{(v-u)\ I(Y_i | [v, \infty)) + (Y_i - u)\ I(Y_i | [u, v))\}\right)} = \lambda . \qquad (4.2)$$

Proof

See Crow and Shimi (1971).

Theorem 4.4

The naive MLE $\hat{\lambda}_{nq}$ is a strongly consistent estimator of λ_q, $q = 0, \ldots, t-1$, as $n \to \infty$.

Proof:

If $\lambda_q = 0$, then $F(S_{q+1}) = 0$ and, hence, $\sum_{r=1}^{n} G_{rq} = 0$ a.s. Thus,

$\hat{\lambda}_{nq} = 0$ a.s.. If $\lambda_q > 0$, then

$$\left(\sum_{r=1}^{n} G_{rq}\right)\Big/\left(\sum_{r=1}^{n} A_{rq}\right) \longrightarrow \frac{E(G_{1q})}{E(A_{1q})} = \lambda_q \text{ a.s. as } n \to \infty,$$

by the strong law of large numbers and (4.2).

4.4 The MLE of the λ_q.

In the next theorem we will find the values of the λ_q, $q = 0,\ldots, t-1$, say $\tilde{\lambda}_{nq}$, which will maximize L under the restriction that $\tilde{\lambda}_{n0} \leq \tilde{\lambda}_{n1} \leq \cdots \leq \tilde{\lambda}_{n(t-1)}$. It will, also, be shown that $\tilde{\lambda}_{nq}$ is a strongly consistent estimator of λ_q.

Theorem 4.5

The MLE $\tilde{\lambda}_{nq}$ of λ_q, $q = 0,\ldots, t-1$, which maximizes $L = \prod_{r=1}^{n} L_r$ under the restriction that $\tilde{\lambda}_{n0} \leq \tilde{\lambda}_{n1} \leq \cdots \leq \tilde{\lambda}_{n(t-1)}$ is given by

$$\tilde{\lambda}_{nq} = \min_{v \geq q} \max_{u \leq q} \frac{\sum_{r=1}^{n} \sum_{d=u}^{v} G_{rd}}{\sum_{r=1}^{n} \sum_{d=u}^{v} A_{rd}} . \tag{4.3}$$

Proof:

From Lemma (4.1)

$$L = \prod_{j=0}^{t-1} \lambda_j^{\sum_{r=1}^{n} G_{rj}} \exp\left(-\lambda_j \sum_{r=1}^{n} A_{rj}\right) .$$

Applying the results of Brunk (1958) yields (4.3).

Remark 4.6

Brunk (1958), page 447, explains a method for determining $\tilde{\lambda}_{nq}$. Let $\hat{\lambda}_{nq}$ be the naive MLE of λ_q given in Corollary 4.2. If $\hat{\lambda}_{n0} \le \hat{\lambda}_{n1} \le \cdots \le \hat{\lambda}_{n(t-1)}$, then $\tilde{\lambda}_{nq} = \hat{\lambda}_{nq}$. If for some i, $\hat{\lambda}_{ni} > \hat{\lambda}_{n(i+1)}$, then replace $\hat{\lambda}_{ni}$ and $\hat{\lambda}_{n(i+1)}$ by

$$[\sum_{r=1}^{n} (G_{ri} + G_{r(i+1)})]/[\sum_{r=1}^{n} (A_{ri} + A_{r(i+1)})].$$

If a reversal still exists, replace by appropriate averages. That is, if

$$[\sum_{r=1}^{n} (G_{ri} + G_{r(i+1)})]/[\sum_{r=1}^{n} (A_{ri} + A_{r(i+1)})] > \hat{\lambda}_{n(i+2)},$$

then replace $\hat{\lambda}_{ni}$, $\hat{\lambda}_{n(i+1)}$, and $\hat{\lambda}_{n(i+2)}$, by

$$[\sum_{r=1}^{n} (G_{ri} + G_{r(i+1)} + G_{r(i+2)})]/[\sum_{r=1}^{n} (A_{ri} + A_{r(i+1)} + A_{r(i+2)})].$$

Continue averaging whenever there is a reversal. This will yield a monotone increasing sequence, $\tilde{\lambda}_{n0} \le \tilde{\lambda}_{n1} \le \tilde{\lambda}_{n1} \le \cdots \le \tilde{\lambda}_{n(t-1)}$, which are the MLE's of the λ_q's subject to $\lambda_0 \le \lambda_1 \le \cdots \le \lambda_{t-1}$.

We now have

Theorem 4.7

The estimates $\tilde{\lambda}_{nq}$ of λ_q, $q = 0,\ldots, t-1$, given in Theorem 3.5 are strongly consistent as $n \to \infty$.

Proof:

This is immediate from Remark 4.6 and the fact that the $\hat{\lambda}_{nq}$'s are strongly consistent estimators.

4.4 Asymptotic Normality of the Estimators.

Without any loss of generality denote by λ_i, $i = 1,2,\ldots,m$, the set of all λ_i, $i = 0,1,\ldots, t-1$ such that $\lambda_i \neq 0$. In the following the asymptotic normality of the vector $(\hat{\lambda}_{n1},\ldots,\hat{\lambda}_{nm})$, suitably normalized, is established. The normalizing sequence is determined from the experimental outcomes and λ_i, $i = 1,\ldots,m$. Furthermore, it is shown that the dispersion matrix of the limiting distribution is the m×m identity matrix I.

Let

$$T_{nj} = \sum_{i=1}^{n} (G_{ij} - E(G_{ij}))/n,\ j = 1,2,\ldots,m, \tag{4.4}$$

and

$$U_{nj} = \sum_{i=1}^{n} (A_{ij} - E(A_{1j}))/n,\ j = 1,2,\ldots,m. \tag{4.5}$$

Define

$$Y_{nj} = \begin{cases} T_{nj} & j = 1,2,\ldots,m \\ U_{n(j-m)} & j = m+1,\ldots,2m \end{cases} \tag{4.6}$$

and $\Sigma = (\sigma_{ij})$ is a 2m x 2m matrix, where

$$\sigma_{ij} = \text{Cov}(Y_{1i}, Y_{1j}), \ i, j = 1,2,\ldots,2m.$$

By the multivariate central limit theorem, it follows that

$$\sqrt{n}\,(Y_{n1}, Y_{n2},\ldots,Y_{n,2m}) \text{ is AN } (\underline{0},\Sigma).$$

We shall use the following theorem given in Rao (1968, page 322), and we state it here for easy reference. "Let $\underline{T}_n$ be a k-dimensional statistic $(T_{1n},\ldots,T_{kn})$ with the asymptotic distribution of $\sqrt{n}\,((T_{1n}-\theta_1),\ldots,(T_{kn}-\theta_k))$ being k-variate normal with mean zero and dispersion matrix $\Sigma = (\sigma_{ij})$. Let $f_1,\ldots,f_q$ be q functions of k variables and each f_i be totally differentiable. Then the asymptotic distribution of $\sqrt{n}\left(f_i(t_{1n},\ldots,T_{kn}) - f_i(\theta_1,\ldots,\theta_k)\right)$, $i = 1,2,\ldots,q$, is q-variate normal with mean zero and dispersion matrix $\Gamma\,\Sigma\,\Gamma'$, where $\Gamma = (\partial f_i/\partial\theta_j)$."

Let $f_i(y_1,y_2,\ldots,y_{2m})$ be the real-valued function of 2m variables defined by

$$f_i(y_1,\ldots,y_{2m}) = y_i - \lambda_i y_{i+m}, \ i = 1,2,\ldots,m. \tag{4.7}$$

Then

$$\partial f_i/\partial y_j = \begin{cases} 1 & \text{if } j = i \\ -\lambda_i & \text{if } j = i + m \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

Therefore, one can show that $\Gamma \Sigma \Gamma' = (\tau_{ij})$, where

$$\tau_{ij} = \mathrm{Var}(G_{1i} - \lambda_i A_{1i}), \text{ if } i = j,$$

and

$$\tau_{ij} = \mathrm{Cov}(G_{1i}, G_{1j}) - \lambda_i \,\mathrm{Cov}(A_{1i}, G_{1j}) - \lambda_j \,\mathrm{Cov}(G_{1i}, A_{1j})$$

$$+ \lambda_i \lambda_j \,\mathrm{Cov}(A_{1i}, A_{1j}), \quad \text{if } i \neq j.$$

Hence, by the theorem mentioned above,

$$\sqrt{n}\left(\sum_{r=1}^{n} \frac{G_{r1} - \lambda_1 A_{r1}}{n}, \ldots, \sum_{r=1}^{n} \frac{G_{rm} - \lambda_m A_{rm}}{n}\right)$$

is AN $(\underline{0}, \Gamma \Sigma \Gamma')$.

Let $A_i = E(A_{1i})$, $i = 1,2,\ldots,m$. Note that $\lambda_i \neq 0$ implies $A_i \neq 0$. One can see then that the asymptotic distribution of

$$\underline{X}_n = \sqrt{n}\left(\sum_{r=1}^{n} \frac{G_{r1} - \lambda_1 A_{r1}}{nA_1}, \ldots, \sum_{r=1}^{n} \frac{G_{rm} - \lambda_m A_{rm}}{nA_n}\right)$$

is $N(\underline{0}, \Sigma_1)$, where $\Sigma_1 = (\tau_{ij}/A_i A_j)$.

Since $\frac{1}{n} \sum_{r=1}^{n} A_{ri} \to A_i$ a.s., one can, also, show that

$$\underline{Y}_n = \sqrt{n}\left(\sum_{r=1}^{n} \frac{G_{r1} - \lambda_1 A_{r1}}{\sum_{r=1}^{n} A_{r1}}, \ldots, \sum_{r=1}^{n} \frac{G_{rm} - \lambda_m A_{rm}}{\sum_{r=1}^{n} A_{rm}}\right)$$

is AN $(\underline{0},\Sigma_1)$. But since

$$\hat{\lambda}_i = \sum_{r=1}^{n} G_{ri} \Big/ \sum_{r=1}^{n} A_{ri}, \quad i = 1,2,\ldots,m,$$

it follows that $\sqrt{n}\,(\hat{\lambda}_{n1}-\lambda_1,\ldots,\hat{\lambda}_{nm}-\lambda_m)$ is AN $(\underline{0},\Sigma_1)$.

Let $D_i = \tau_{ij}/A_i^2$, $i = 1,2,\ldots,m$. Then

$$\sqrt{n}\left(\frac{\hat{\lambda}_{n1}-\lambda_1}{\sqrt{D_1}},\ldots,\frac{\hat{\lambda}_{nm}-\lambda_m}{\sqrt{D_m}}\right)$$

is AN $(\underline{0},\Sigma_2)$, where

$$\Sigma_2 = (\delta_{ij}) = (\tau_{ij}/A_iA_j(D_iD_j)^{1/2}). \tag{4.9}$$

Note that $\delta_{ij} = 1$ if $i = j$, and

$$D_i = \frac{1}{A_i^2}\ \mathrm{Var}(G_{1i}-\lambda_iA_{1i}),\ i = 1,2,\ldots,m. \tag{4.10}$$

Observe, also, that $E(G_{1i}) = \lambda_iE(A_{1i})$, $i=1,2,\ldots,m$, by Equation (4.2). Now, assuming that $\lambda_i \neq 0$ and $\lambda_s \neq 0$, this implies

$$0 = \frac{\partial}{\partial\lambda_s}\sum_{k=1}^{\infty}\int\!\!\ldots\!\int\left\{\frac{1}{\lambda_i}\cdot G_i(x_1,\ldots,x_{k-1})\right.$$

$$\left. - A_i(x_1,\ldots,x_{k-1},\ T-\sum_{\ell=1}^{k-1}x_\ell)\right\}\prod_{j=0}^{t-1}\lambda_j^{G_j}\exp(-\lambda_jA_j)\prod_{i=1}^{k-1}dx_i.$$

Thus, if i = s,

$$0 = \frac{1}{\lambda_i^2} E(G_{1i})^2 - \frac{1}{\lambda_i^2} E(G_{1i}) - \frac{2}{\lambda_i} E(G_{1i}-A_{1i}) + E(A_{1i})^2 .$$

Using (4.2) again gives

$$E(G_{1i}) = Var(G_{1i}) + \lambda_i^2 Var(A_{1i}) - 2\lambda_i Cov(G_{1i},A_{1i})$$

$$= Var(G_{1i}-\lambda_i A_{1i}) . \tag{4.11}$$

If i ≠ s, then

$$0 = \sum_{k=1}^{\infty} \int\cdots\int \left\{\frac{1}{\lambda_i} G_i - A_i\right\} \left\{\frac{1}{\lambda_s} G_s - A_s\right\}$$

$$\prod_{j=0}^{t-1} \lambda_j^{G_j} \exp(-\lambda_j A_j) \prod_{i=1}^{k-1} dx_i ,$$

and this gives

$$0 = E(G_{1i}G_{1s}) - \lambda_s E(G_{1i}A_{1s}) - \lambda_i E(A_{1i}G_{1s})$$
$$+ \lambda_i\lambda_s E(A_{1i}A_{1s}) .$$

Hence,

$$\lambda_s E(G_{1i})E(A_{1s}) + \lambda_i E(A_{1i})E(G_{1s}) - E(G_{1i})E(G_{1s})$$

$$- \lambda_i\lambda_s E(A_{1i})E(A_{1s}) = Cov(G_{1i},G_{1s}) - \lambda_s Cov(G_{1i},A_{1s})$$

$$- \lambda_i Cov(A_{1i},G_{1s}) + \lambda_i\lambda_s Cov(A_{1i},A_{1s}) . \tag{4.12}$$

Since $E(G_{1\ell}) = \lambda_\ell E(A_{1\ell})$, it follows that the left-hand side of (4.12) is equal to zero and, hence, $\Sigma_2 = I$. Thus,

$$\sqrt{n}\left(\frac{\hat{\lambda}_{n1}-\lambda_1}{\sqrt{D_1}}, \ldots, \frac{\hat{\lambda}_{nm}-\lambda_m}{\sqrt{D_m}}\right)$$

is AN $(\underline{0}, I)$. Note that if $\lambda_1 < \ldots < \lambda_m$ then $\hat{\lambda}_{ni} = \hat{\lambda}_i$ a.s. for $n \geq n_0$, say. Thus,

$$\sqrt{n}\left(\frac{\tilde{\lambda}_{n1}-\lambda_1}{\sqrt{D_1}}, \ldots, \frac{\tilde{\lambda}_{nm}-\lambda_m}{\sqrt{D_m}}\right)$$

is AN $(\underline{0}, I)$. By (4.11)

$$D_i = \frac{1}{A_i^2} \operatorname{Var}(G_{1i}-\lambda_i A_{1i}) = \frac{E(G_{1i})}{A_i^2} .$$

Using (4.2), this gives

$$D_i = \frac{\lambda_i^2}{E(G_{1i})} .$$

Note that since $\frac{1}{n}\sum_{r=1}^{n} G_{ri} \to E(G_{1i})$ a.s., then $\dfrac{\lambda_i^2}{\frac{1}{n}\sum_{r=1}^{n} G_{ri}} \to D_i$ a.s.

The next result is immediate.

<u>Theorem 4.8</u>

The asymptotic distribution of

$$\left[(\hat{\lambda}_{n1}-\lambda_1)\Big/\sqrt{\frac{\lambda_1^2}{\sum_{r=1}^{n} G_{r1}}}, \ldots, (\hat{\lambda}_{nm}-\lambda_m)\Big/\sqrt{\frac{\lambda_m^2}{\sum_{r=1}^{n} G_{rm}}}\right] \text{ is } N(\underline{0}, I) \text{ as } n \to \infty.$$

As before, if $\lambda_1 < \dots < \lambda_m$, then the same result holds if the $\hat{\lambda}_{ni}$ are replaced by $\tilde{\lambda}_{ni}$.

Remark 4.9

If one assumes that $\lambda_0 \geq \lambda_1 \geq \dots \geq \lambda_{t-1}$ then the likelihood $L = \prod_{r=1}^{n} L_r$ is the same and hence the naive MLE $\hat{\lambda}_{nq}$ of λ_q is the same and is consistent. The MLE $\tilde{\lambda}_{nq}$ may be easily found by applying again the results of Brunk (1958). Of course $\tilde{\lambda}_{nq}$ is also consistent in this case.

The reader is probably aware, at this point, that if t=1 then this class of distribution reduces to the class considered in the last section when $g(x) = x$. The MLE is equivalent in both situations.

5. NONPARAMETRIC ESTIMATION

5.1 Introduction.

The concept of "failure rate" is a very important practical concept in reliability and has motivated several very useful classes of distributions, e.g., Increasing Failure Rate (IFR) class, Decreasing Failure Rate (DFR) class, U-Shaped Failure Rate class. The failure rate $r(\cdot)$ of a distribution function F having derivative f is defined by

$$r(x) = f(x)/[1 - F(x)] \qquad \text{for } F(x) < 1$$

and

$$r(x) = \infty \qquad \text{for } F(x) = 1.$$

The estimation problem that we shall be concerned with in this section can be summarized in the following way. The life-time distribution of the items to be tested is assumed to have an increasing failure rate over the interval $[0,T)$, i.e., IFR on $[0,T)$. No other assumptions about the distribution or its failure rate is given outside that interval. The assumption of increasing failure rate can be changed to decreasing failure rate and the same results will follow with the obvious modifications. Data are assumed to arise from Testing Plan A.

Let $[\alpha_F, \beta_F]$ be the support of the c.d.f. F. The notion of IFR on $[0,T)$ is made more precise by the following definition.

Definition 5.1

Let T be a fixed positive real number. A c.d.f. F, $F(0) = 0$, is said to be IFR (Increasing Failure Rate) on $[0,T)$ iff it satisfies one of the following conditions: (i) $-\log[1-F(x)]$ is convex on $[\alpha_F, \beta_F]$, $0 \le \alpha_F \le \beta_F \le T$ and $F(\beta_F) = 1$ if $\beta_F < T$; or (ii) the part of the support of F in $[0,T]$ is empty.

Let $\mathcal{F} = \{F: F \text{ is IFR on } [0,T)\}$.

The following theorem is similar to a theorem concerning IFR distribution given by Marshall and Proschan (1965), and we shall omit its proof because of this similarity.

Theorem 5.2

Suppose $F \in \mathcal{F}$, $0 < Z \le \beta_F$. Then F is absolutely continuous on $[0,Z)$. Note that F may take a jump at β_F if $\beta_F < T$.

Using the definition of failure rate and the above theorem one can very easily prove the following.

Theorem 5.3

(i) $F \in \mathcal{F}$ iff $r(\cdot)$ is nondecreasing on $[0,\beta_F)$, $0 \le \alpha_F \le \beta_F \le T$, and $F(\beta_F) = 1$ if $\beta_F < T$.

(ii) The part of the support of F in $[0,T]$ is empty iff $r(x) = 0$ on $[0,T]$.

(iii) If $F \in \mathcal{F}$, then for $x \in [0, \beta_F)$

$F(x) = 1 - \exp(-R(x))$, and

$f(x) = r(x)\exp(-R(x))$, where

$$R(x) = \int_0^x r(y)dy.$$

The class $\mathcal{F}$ includes the usual class of IFR distributions. It is easy to show that there exists no sigma-finite measure relative to which all the distributions in $\mathcal{F}$ are absolutely continuous.

Since we are dealing with a nonparametric family of distributions for which there exists no sigma-finite measure relative to which all the measures induced by $\mathcal{F}$ are absolutely continuous, the usual concept of maximum likelihood estimation cannot be applied. The general definition of MLE due to Kiefer and Wolfowitz (1956) is used in this section to determine the MLE of the life-time distribution F over $[0,T)$, where $F \in \mathcal{F}$ and data arise from Testing Plan A. It is also shown that this MLE is strongly consistent as n, the number of original items, tends to infinity.

In this section let $d(n)$ denote the total number of distinct failures in $[0,T)$ in the combined n experiments. Recalling the notation given in Section 2, observe that

$$0 \leq d(n) \leq \sum_{r=1}^{n} (K_r - 1).$$

Also, let $0 = Z_0 < Z_1 < \dots < Z_{d(n)}$ be the ordered, distinct, failure times X_{jr}, $j = 1, \dots, K_r-1$, $r = 1, \dots, n$. Finally, let $p(n)$ be the number of $Y_{K_r r}$'s, $r = 1, \dots, n$, strictly greater than $Z_{d(n)}$.

5.2 Maximum Likelihood Estimate.

In this section we shall find the MLE of that part of a life-time distribution $F \in \mathcal{F}$ over the interval $[0,T)$ when data arise from

Testing Plan A. The following general definition of a maximum likelihood estimate is due to Kiefer and Wolfowitz (1956) and is needed to determine the MLE of $F \in \mathcal{F}$ for the two reasons mentioned earlier.

Definition 5.4

Let Ω be a sample space, $\mathcal{B}$ a σ-field on Ω, $\mathcal{P}$ a family of probability measures on $\mathcal{B}$ and Θ a set indexing the elements of $\mathcal{P}$ by $P(\cdot|\theta)$, $\theta \in \Theta$. Let $\underset{\sim}{X}$ be a random vector defined on Ω with distribution function determined by $P(\cdot|\theta_0)$, $\theta_0 \in \Theta$. If $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ denotes a random sample from $P(\cdot|\theta_0)$ then the MLE of θ_0 is $\hat{\theta}$ if $\hat{\theta} \in \Theta$ and

$$\sup_{\theta \in \Theta} \frac{\prod_{r=1}^{n} \dfrac{dP(\underset{\sim}{X}_r|\theta)}{d(P(\underset{\sim}{X}_r|\theta) + P(\underset{\sim}{X}_r|\hat{\theta}))}}{\prod_{r=1}^{n} \left\{1 - \dfrac{dP(\underset{\sim}{X}_r|\theta)}{d(P(\underset{\sim}{X}_r|\theta) + P(\underset{\sim}{X}_r|\hat{\theta}))}\right\}} = 1$$

where

$$\frac{dP(\cdot|\theta_1)}{d(P(\cdot|\theta_1) + P(\cdot|\theta_2))}$$

denotes the Radon-Nikodym derivative of $P(\cdot|\theta_1)$ with respect to $P(\cdot|\theta_1) + P(\cdot|\theta_2)$.

The Kiefer and Wolfowitz concept of MLE will now be considered within the framework of Testing Plan A and for life-time distributions $F \in \mathcal{F}$. Let $\Omega = \{\phi$, and all finite sequences of non-negative numbers whose sum is less than $T\}$, where ϕ is the empty set. Also, let $X_i = \{\underset{\sim}{x} \in \Omega$ which have exactly i elements$\}$, $i = 0, 1, \ldots$. Then $\Omega = \bigcup_{i=0}^{\infty} X_i$. We

define a set A to be measurable in Ω if and only if $A = \sum_{i=0}^{\infty} A_i$ and A_i is Borel measurable in X_i. Let $\mathcal{B}$ the σ-field of measurable sets in Ω.

For each $F \in \mathcal{F}$ we will define a probability measure $P(\cdot|F)$ on $\mathcal{B}$ and will denote the collection of all such measures by $\mathcal{P}$. These probability measures will be defined first on the Borel measurable sets of each X_i. Some preliminary notation is needed. Denote by $\lambda(\cdot|i, F)$ the product measure on R^i (Eucidean i-th space) induced by F, where $\lambda(\cdot|0, F)$ is defined to be one. Also, recall that $F(x^-) = \lim_{\varepsilon \to 0} F(x-\varepsilon)$, $\varepsilon > 0$, and products of the form $\prod_{j=1}^{0}$ and sums of the form $\sum_{j=1}^{0}$ are 1 and 0, respectively. For each Borel measurable set $A_i \subseteq X_i$ and $F \in \mathcal{F}$ define the measure $P(\cdot|F)$ to be

$$P(A_i|F) = \int_{A_i} \{1 - F([T-\sum_{j=1}^{i} x_j]^-)\} \, d\lambda(\underset{\sim}{x}|i,F).$$

For any $A \in \mathcal{B}$ we define $P(A|F)$ to be

$$P(A|F) = \sum_{i=0}^{\infty} P(A_i|F)$$

where $A_i = A \cap X_i$.

This definition is motivated by the integrand of equation (2.7).

Note that for each $F \in \mathcal{F}$

$$P(\Omega|F) = P(\sum_{i=0}^{\infty} X_i|F) = \sum_{i=0}^{\infty} P(X_i|F)$$

$$= \sum_{i=0}^{\infty} \text{Prob}(K = i + 1|F) = 1.$$

Thus, for each $F \in \mathcal{F}$, $P(\cdot|F)$ is a probability measure on $\mathcal{B}$

The Kiefer and Wolfowitz concept of maximum likelihood estimate together with our definition of the measures $P(\cdot|F) \in \mathcal{P}$, $F \in \mathcal{F}$, yields the MLE $\tilde{F}_n$ of F on $[0,T)$ described in the next theorem.

Let $I(\cdot|S)$ be the indicator function of S. Also, let $n_r(y)$ denote

$$\sum_{i=1}^{k_r} I(Y_{ir}|[y,\infty)).$$

Theorem 5.5

The MLE $\tilde{F}_n$ of F has failure rate $\tilde{r}_n$ where $\tilde{r}_n$ is constant over $[Z_q, Z_{q+1})$, $q = 0,\ldots,d(n)$, and

$$\tilde{r}_n(Z_q) = \min_{d(n)+1 \geq v \geq q+1} \max_{0 \leq u \leq q} \frac{\sum_{r=1}^{n} \sum_{j=1}^{K_r - 1} I(X_{jr}|[Z_n, Z_v))}{\sum_{r=1}^{n} \int_{Z_u}^{Z_v} n_r(y)dy}. \tag{5.3}$$

Proof:

The proof of this theorem follows in a straightforward manner from the Kiefer-Wolfowitz definition of MLE using the probability measures we introduced above and Brunk's (1958) results.

Remark 5.6

We will now give a useful method for determining $\tilde{r}_n$. Let T_{qn} be the time on test over $[Z_q, Z_{q+1})$ (i.e. $T_{qn} = \sum_{r=1}^{n} \int_{Z_q}^{Z_{q+1}} n_r(y)dy$), $q = 0,\ldots, d(n)$. If $(T_{0n})^{-1} \leq (T_{1n})^{-1} \leq \ldots \leq (T_{d(n)n})^{-1}$ then $\tilde{r}_n(Z_q) = (T_{qn})^{-1}$, $q = 0,\ldots,d(n)$. If for some i, $(T_{in})^{-1} > (T_{(i+1)n})^{-1}$ then replace $(T_{in})^{-1}$ and $(T_{(i+1)n})^{-1}$ by $2(T_{in} + T_{(i+1)n})^{-1}$.

If a reversal still exists, replace by appropriate averages. That is, if $2(T_{in} + T_{(i+1)n})^{-1} > (T_{(i+2)n})$, then replace $(T_{in})^{-1}$, $(T_{(i+1)n})^{-1}$ and

$(T_{(i+2)n})^{-1}$ by $3(T_{in} + T_{(i+1)n} + T_{(i+2)n})^{-1}$.

Continue averaging whenever there is a reversal. This will yield the monotone increasing sequence $\tilde{r}(Z_0) \leq \tilde{r}_n(Z_1) \leq \cdots \leq \tilde{r}_n(Z_{d(n)})$ given by (5.3).

5.3 Strong Consistency of $\tilde{F}_n$.

The main result of this section is that the MLE of F on $[0,T)$ converges uniformly a.s. to F as the number of items put on test at time zero increases. To accomplish this we will prove a convergence theorem for $\tilde{r}_n(x)$, $x\epsilon[0,T)$, $\tilde{r}_n(x)$ defined in Theorem 5.5. This result will allow us to easily prove the main result plus several corollaries. Furthermore, since the failure rate of a life-time distribution is an important practical concept, the convergence theorem for $\tilde{r}_n(x)$ is also a significant practical result.

We will need several theorems before we can prove the convergence theorem for $\tilde{r}_n(x)$. The next theorem involves rewriting $\tilde{r}_n(x)$, given in Theorem 5.5, in a form we need to show consistency.

Let $R(u,v)$ denote $\sum_{r=1}^{n} \sum_{j=1}^{K_r-1} I(X_{jr}|[u,v))$ and $S(u,v)$ denote $\sum_{r=1}^{n} \int_{n}^{v} n_r(y)dy$.

Theorem 5.7

Let $x\epsilon[0,T)$ and

$$Z_n(x) = \max_{0\leq i\leq d(n)} \{Z_i | Z_i \leq x\}.$$

Then

$$\tilde{r}_n(x) = \inf_{x \le v < T} \sup_{u < Z_n(x)} \frac{R(u,v)}{S(u,v)} . \tag{5.4}$$

Proof:

Follows directly from Theorem 5.5.

To show consistency of $\tilde{r}_n$ we need the next two theorems. Let

$$M_n(u,v) = \frac{\sum_{r=1}^{n} \sum_{j=1}^{K_r - 1} I(X_{jr} | [u,v))}{\sum_{r=1}^{n} \int_u^v n_r(y)dy}, \quad 0 < u < v < T,$$

and let I_F be the intersection of the support of F with $[0,T]$.

Theorem 5.8

Let $0 \le u_0 \le v_0 \le T$ be fixed where, $0 \le u_0 < T$ if $I_F = \phi$, $0 \le u_0 < v_0 < \beta_F$ if $I_F = [\alpha_F, \beta_F]$. Then, as $n \to \infty$

i) $M_n(u_0,v)$ converges uniformly, a.s., in $v_0 \le v \le T$;

ii) $M_n(u,v_0)$ converges uniformly, a.s. in $0 \le u \le u_0$.

Proof:

Let X_1, $X_2, \ldots$, be a sequence of independent, identically distributed (i.i.d.) random variables with c.d.f. F, $F(0) = 0$. Let N_1 be the first integer such that $\sum_{i=1}^{N_1} X_i \ge T$, N_2 the first integer such that $\sum_{i=N_1+1}^{N_2} X_i \ge T$, N_3 the first integer such that $\sum_{i=N_1+N_2+1}^{N_3} X_i \ge T$, and so on. Then $N_1, N_2, \ldots$, is a sequence of i.i.d. random variables. Using the Glivenko-Cantelli theorem one may show that as $n \to \infty$

$$B_n(u,v) \xrightarrow{U} F(v^-) - F(u^-), \text{ a.s.}, -\infty < u < v < \infty, \tag{5.5}$$

where

$$B_n(u,v) = \frac{\sum_{i=1}^{N(n)} I(X_i \,|\, [u,v))}{N(n)} , \quad N(n) = \sum_{r=1}^{n} N_r .$$

Also, using the strong law of large numbers and the Glivenko-Cantelli theorem it is easy to show that as $n \to \infty$

$$C_n(u,v) \xrightarrow{U} \frac{1}{E(K_1)} [F_1(v^-) - F_1(u^-)] \text{ a.s.} \tag{5.6}$$

for $-\infty < u < v < \infty$, where

$$C_n(u,v) = \frac{\sum_{r=1}^{n} I(X_{K_r,r} \,|\, [u,v))}{A(n)} , \quad A(n) = \sum_{r=1}^{n} K_r ,$$

and F_1 is the c.d.f. of $X_{K_1,1}$. We may conclude from (5.5) and (5.6) that as $n \to \infty$

$$D_n(u,v) = B_n(u,v) - C_n(u,v) \tag{5.7}$$

converges uniformly, a.s. for $-\infty < u < v < \infty$.

Similarly one may show that as $n \to \infty$

$$S_n(u,v) \text{ converges uniformly a.s. on } 0 \le u < v \le T \tag{5.8}$$

where

$$S_n(u,v) = \frac{\sum_{r=1}^{n} \int_u^v n_r(y)dy}{A(n)} .$$

Also, observe that for n_0 sufficiently large,

$$(S_n(u_0,v))^{-1} \text{ is uniformly bounded, a.s.,} \tag{5.9}$$

on $v_0 \leq v < T$, $n \geq n_0$,

$$D_n(u_0,v)) \text{ is uniformly bounded, a.s.} \tag{5.10}$$

on $v_0 \leq v < T$, $n \geq n_0$

$$(S_n(u,v_0))^{-1} \text{ is uniformly bounded, a.s.} \tag{5.11}$$

on $0 \leq u \leq u_0$, $n \geq n_0$, and

$$D_n(u,v_0) \text{ is uniformly bounded, a.s.} \tag{5.12}$$

on $0 \leq u \leq u_0$, $n \geq n_0$.

The proof is completed since (5.7) - (5.10) imply (i) and (5.7), (5.8), (5.11) and (5.12) imply (ii).

<u>Theorem 5.9</u>

Let F be IFR on $[0,T)$ with failure rate r on $[0,T)$. Then, for $0 \leq u < v < T$ fixed, where $0 \leq u < \beta$ if $I_F = [\alpha,\beta]$,

$$r(u) \leq \frac{E(\sum_{i=1}^{K-1} I(X_i|[u,v)))}{E(\int_u^v n(y)dy)} \leq r(v) \tag{5.13}$$

where $K = K_1$, $n(\cdot) = n_1(\cdot)$ and $X_i = X_{i1}$, $i = 1,2,\ldots$.

Proof:

If $I_F = \phi$ then F has failure rate 0 on $[0,T)$ and, thus, (5.13) follows. If $I_F = \{\beta\}$ then $F(\beta) = 1$. Consequently, $r(x) = \infty$, $x \geq \beta$, and $r(x) = 0$, $x < \beta$. Also,

$$E(\sum_{i=1}^{K-1} I(X_i|[u,v))) = 0 \text{ for } u < v \leq \beta$$

and

$$E(\sum_{i=1}^{K-1} I(X_i|[u,v))) \geq 1 \text{ for } u < \beta < v.$$

Further,

$$E(\int_u^v n(y)dy) > 0 \text{ for } u < \beta.$$

Thus, (5.13) easily follows when $I_F = \{\beta\}$.

Now, assume $I_F = [\alpha,\beta]$, $0 \leq \alpha < \beta \leq T$. Also, recall that by Theorem 4.3, we know that if H is a c.d.f. with failure rate constant, say, λ, on $[a,b)$, then

$$\frac{E_H(\sum_{i=1}^{K-1} I(X_i|[a,b)))}{E_H(\int_a^b n(y)dy)} = \lambda. \tag{5.14}$$

Case 1.

$$0 < u < v < \beta.$$

If F has a nondecreasing step-function failure rate on $[u,v)$ then (5.13) holds by a simple application of (5.14). To prove that

(5.13) holds in general for this case, let $r_n(x)$, $n = 1,2,\dots$, $x\varepsilon[0,\beta)$ be a sequence of real valued functions such that $r_n(x) = r(x)$, $x\varepsilon[0,u)$, $r_n(x)$ is a nondecreasing step-function on $[u,v)$ and $r_n(x) \uparrow r(x)$ on $[u,v)$. Note that $r_n(x) \leq r(x) \leq r(v) < \infty$. Thus, by the Lebesgue Dominated Convergence theorem, as $n \to \infty$

$$\int_0^y r_n(x)dx \to \int_0^y r(x)dx, \; y \; \varepsilon \; [0,v).$$

Therefore,

$$F_n(y) \equiv 1 - \exp\{-\int_0^y r_n(x)dx\} \to 1 - \exp\{-\int_0^y r(x)dx\}$$

$$= F(y), \; y \; \varepsilon \; [0,v), \text{ as } n \to \infty.$$

Let $F_n(y) = F(y)$, $y \geq v$. Then F_n, $n = 1,2,\dots$, is absolutely continuous on $[0,v)$, continuous from the right on $[v,\infty)$, since F is, and $F_n(0) = 0$, $F_n(\infty) = 1$. Thus, F_n is a sequence of distribution functions, $F_n(y) \to F(y)$, $y \; \varepsilon \; (-\infty,\infty)$, as $n \to \infty$. Let

$$S_k = \{\sum_{i=1}^{k-1} x_i < T, \; \sum_{i=1}^{k} x_i \geq T\}, \; k = 1,2,\dots .$$

By the Helly-Bray theorem (Loeve (1963))

$$P[K=k|F_n] = \int\cdots\int_{S_k} \prod_{i=1}^{k} dF_n(x_i) \to \int\cdots\int_{S_k} \prod_{i=1}^{k} dF(x_i)$$

$$= P[K=k|F], \; n \to \infty, \; k = 1,2,\dots . \tag{5.15}$$

Let $p_n(k) = P[K=k|F_n]$, $n = 1,2,\ldots$, $k = 1,2,\ldots$, and $p(k) = P[K=k|F]$, $k = 1,2,\ldots$. By Rao ((1968), page 106) and (5.15)

$$\sum_{k=0}^{\infty} |p_n(k) - p(k)| \to 0, \ n \to \infty. \tag{5.16}$$

Now, note that

$$\sum_{i=1}^{K-1} I(X_i|[u,v)) \leq [\tfrac{T}{u}], \text{ a.s., } u > 0 \tag{5.17}$$

and

$$n(y) \leq [\tfrac{T}{y}], \text{ a.s., } y > 0 \tag{5.18}$$

where $[x]$ denotes the largest integer less than or equal to x.

Thus, since $u > 0$ and (5.16) holds

$$|E_{F_n}(\sum_{i=1}^{K-1} I(X_i|[u,v))) - E_F(\sum_{i=1}^{K-1} I(X_i|[u,v)))|$$

$$= |\sum_{k=1}^{\infty} \int_{A_k} \sum_{i=1}^{k-1} I(x_i|[u,v)) \prod_{j=1}^{k} dF_n(x_j)$$

$$- \sum_{k=1}^{\infty} \int_{A_k} \sum_{i=1}^{k-1} I(x_i|[u,v)) \prod_{j=1}^{k} dF(x_j)|$$

$$\leq [\tfrac{T}{u}] \sum_{k=1}^{\infty} |p_n(k) - p(k)| \to 0, \text{ as } n \to \infty.$$

Also,

$$\int_u^v n(y)dy \leq (v-u)\, n(u) \leq (v-u)[\tfrac{T}{u}]$$

by (5.16). Hence,

$$E_{F_n}\left(\sum_{i=1}^{K-1} I(X_i|[u,v))\right) \to E_F\left(\sum_{i=1}^{K-1} I(X_i|[u,v))\right) \tag{5.19}$$

and

$$E_{F_n}\left(\int_u^v n(y)dy\right) \to E_F\left(\int_u^v n(y)dy\right). \tag{5.20}$$

However, from (5.14) it follows that

$$r_n(y) \le \frac{E_{F_n}\left(\sum_{i=1}^{K-1} I(X_i|[u,v))\right)}{E_{F_n}\left(\int_u^v n(y)dy\right)} \le r_n(v).$$

Taking limits, and using (5.19) and (5.20) gives (5.13).

<u>Case 2</u>.

$$0 = u < v < \beta.$$

Inequalities (5.13) follows easily using the results of Case 1.

Also it is straightforward to use the results of Case 1 to prove (5.13) for

<u>Case 3</u>.

$$\beta < T, \ \beta \le v < T.$$

We now give the convergence theorem for the estimate $\tilde{r}_n(x)$ of $r(x)$, $0 \le x < T$.

Theorem 5.10

Let F be IFR on $[0,T)$ with failure rate r on $[0,T)$. Then,

$$r(x_0^-) \leq \liminf \tilde{r}_n(x_0) \leq \limsup \tilde{r}_n(x_0) \leq r(x_0^+) \text{ a.s.}$$

for each $x_0 \in (0,T)$.

Proof:

Case 1.

$$I_F = \phi.$$

In this case $\tilde{r}_n(x) = 0$ a.s. for $0 \leq x < T$. Since $r(x) = 0$, $0 \leq x < T$, the result follows.

Case 2.

$$I_F = [\alpha,\beta].$$

Let $Z_n(x_0) = \max\limits_{0\leq i\leq d(n)} \{Z_i | Z_i \leq x_0\}$. We will show the right-hand inequality first.

If $\beta < T$ and $\beta \leq x_0 < T$, then $r(x_0^+) = \infty$, since $F(\beta) = 1$. Hence, assume $0 < x_0 < \beta \leq T$. Choose v_0, $x_0 < v_0 < \beta$. Then

$$\tilde{r}_n(x_0) = \inf_{x_0\leq v} \sup_{u<Z_n(x_0)} M_n(u,v)$$

$$\leq \sup_{u<Z_n(x_0)} M_n(u,v_0). \tag{5.21}$$

Let

$$M(a,b) = \frac{E(\sum_{j=1}^{K_1-1} I(X_{j1}|[a,b))}{E(\int_a^b n_1(y)dy)}.$$

Since $0 < x_0 < v_0$, we may apply Theorem 5.8 (ii) and conclude that as $n \to \infty$ $M_n(u,v_0)$ converges uniformly a.s. for $0 \leq u \leq x_0$. Thus, for arbitrary $\varepsilon > 0$ and $n \geq N(\varepsilon)$, say,

$$\tilde{r}_n(x_0) \leq \sup_{u<Z_n(x_0)} (M(u,v_0) + \varepsilon).$$

Since $u < \beta$ we may apply Theorem 5.9 and conclude that $\limsup \tilde{r}_n(x_0) \leq r(v_0) + \varepsilon$. This gives $\limsup \tilde{r}_n(x_0) \leq r(x_0^+)$ a.s. since $x_0 < v_0$ and the right-hand limits exist.

We will now show the left-hand inequality.

Case 2a.

$$0 < \alpha \text{ and } x_0 \in (0,\alpha].$$

Since $r(x_0^-) = 0$ the left-hand inequality holds.

Case 2b.

$$\beta < T \text{ and } \beta \leq x_0 < T.$$

If F takes a jump at β then with probability one $Z_n(x_0) = \beta$ for $n \geq N$, N sufficiently large. But this implies that $\tilde{r}_n(x) = \tilde{r}_n(\beta) = \infty$ for $n \geq N$, $\beta \leq x < T$. Thus, $\liminf \tilde{r}_n(x_0) = \infty$ and, hence, $\liminf \tilde{r}_n(x_0) \geq r(x_0^-)$.

If F does not take a jump at β then $r(\beta^-) = \infty$ and therefore as $n \to \infty$ $Z_n(x_0) \to \beta$ a.s.. Choose u_0, $0 < u_0 < \infty$. Then for N sufficiently large, $u_0 < Z_n(x_0) < \beta$, for $n \geq N$, and, thus,

$$\begin{aligned}\tilde{r}_n(x_0) &= \inf_{x_0 \leq v} \sup_{u < Z_n(x_0)} M_n(u,v) \\ &\geq \inf_{x_0 \leq v} M_n(u_0,v) \\ &\geq \inf_{v_0 \leq v} M_n(u_0,v) \text{ for } u_0 < v_o < \beta.\end{aligned}$$

Apply Theorem 5.8 (i) and conclude that for arbitrary $\varepsilon > 0$, $N(\varepsilon)$ sufficiently large,

$$\tilde{r}_n(x_0) \geq \inf_{v_0 \leq v} (M(u_0,v)-\varepsilon) \text{ a.s., } n \geq N(\varepsilon).$$

By Theorem 5.9, $\tilde{r}_n(x_0) \geq r(u_0) - \varepsilon$, $n \geq N(\varepsilon)$. This gives $\liminf \tilde{r}_n(x_0) \geq r(u_0)$ a.s. for all $u_0 < \beta$. Letting $u_0 \to \beta^-$ gives $\liminf \tilde{r}_n(x_0) = \infty$ a.s.. Since $\tilde{r}_n(x_0^-) = \infty$ for $x_0 \geq \beta$, we have the desired result for Case 2b.

Case 2c.

$$\alpha < x_0 < \beta \leq T.$$

Choose u_0, $\alpha < u_0 < x_0$. Then for N large enough so that $u_0 < Z_n(x_0) < x_0$,

$$\tilde{r}_n(x_0) = \inf_{x_0 \leq v} \sup_{u < Z_n(x_0)} M_n(u,v)$$

$$\geq \inf_{x_0 \leq v} M_n(x_0,v).$$

Applying Theorems 5.8 and 5.9 in the usual manner gives $\liminf \tilde{r}_n(x_0) \geq r(u_0)$ a.s. for all $\alpha < u_0 < x_0$. The result follows.

This completes the proof.

The main result of this section is

Theorem 5.11

Let F be IFR on $[0,T)$ with failure rate r on $[0,T)$. Then $\tilde{F}_n(t) \to F(t)$ uniformly a.s. in $t \in [0,T)$, where

$$\tilde{F}_n(t) = 1 - \exp(-\int_0^t \tilde{r}_n(y)dy).$$

Proof:

Let I_F be the support of F on $[0,T]$. If $I_F = \phi$ the conclusion is clearly true. Note, also, that $\tilde{F}_n(0) = 0$ a.s.. Suppose then that $I_F = [\alpha,\beta]$. By Theorem 5.10 $\tilde{r}_n(t) \to r(t)$, $t \in [0,\beta)$ except possibly on a set of Lebesque measure zero. Let $t \in [0,\beta)$ and let $t \leq t_0 < \beta$ be a continuity point of r. For arbitrary $\varepsilon > 0$ and $N = N(t_0,\varepsilon)$ sufficiently large, $\tilde{r}_n(x) \leq \tilde{r}_n(t_0) \leq r(t_0) + \varepsilon$ for $x \in [0,t]$, $n \geq N$. Thus, by the Lebesque Dominated Convergence theorem

$$\int_0^t \tilde{r}_n(z)dz \to \int_0^t r(z)\, dz \text{ a.s.} \tag{5.22}$$

Since

$$F(t) = 1 - \exp(-\int_0^t r(z)dz), \ t \in [0,\beta),$$

(5.22) implies that

$$\tilde{F}_n(t) \to F(t) \text{ a.s. } t \in [0,\beta). \tag{5.23}$$

Case 1.

F is continuous on $[0,T)$.

Since Testing Plan A and all related random variables are uneffected by the behavior of F on $[T,\infty)$, we may assume without any loss of generality that F is continuous on $[0,T]$.

Case 1a.

$F(T) = 1, \ \beta = T.$

Extend $\tilde{F}_n$ to $(-\infty,\infty)$ by defining $\tilde{F}_n(x) = 0$, $x < 0$, $\tilde{F}_n(x) = 1$, $x \geq T$. Then, by (5.23) as $n \to \infty$

$$\tilde{F}_n(t) \to F(t) \text{ a.s. } t \in (-\infty,\infty). \tag{5.24}$$

Note that $\tilde{F}_n$ is a distribution function. Since F is continuous on $(-\infty,\infty)$, $\tilde{F}_n(t) \to F(t)$ uniformly a.s. $t \in (-\infty,\infty)$, as $n \to \infty$ by Polya's theorem (Eisen, 1969).

Case 1b.

$F(T) < 1, \ \beta = T.$

Since $1-F(T^-) > 0$, there exists a.s. some $m = 1,2,\ldots,$ such that $K_m = 1$. This implies that either $d(n) = 0$, or $p(n) > 1$ and $d(n) \geq 1$,

for $n \geq m$. In any case, $\tilde{r}_n(x) < \infty$ a.s. $x \in [0,T)$, for $n \geq m$. Therefore, $\tilde{F}_n(x) < 1$, $x \in [0,T)$, and hence, $\tilde{F}_n$ may be extended to $[0,T]$ in a continuous manner, for $n \geq m$. Since F is continuous, $\tilde{F}_n(T) \to F(T)$. Let

$$G_n(x) = \begin{cases} \dfrac{\tilde{F}_n(x)}{\tilde{F}_n(T)} & x \in [0,T] \\ 1 & x \in (T,\infty) \\ 0 & x \in (-\infty,0) \end{cases} \tag{5.25}$$

for $n \geq m$, and let

$$G(x) = \begin{cases} \dfrac{F(x)}{F(T)} & x \in [0,T] \\ 1 & x \in (T,\infty) \\ 0 & x \in (-\infty,0)\ . \end{cases} \tag{5.26}$$

Then G_n, $n = m, m + 1,\ldots$, G are distribution functions, G is continuous and $G_n(x) \to G(x)$ a.s. $x \in (-\infty,\infty)$. By Polya's theorem the convergence is uniform. Since $\tilde{F}_n(T)$ is bounded for $n = m, m + 1,\ldots$, this implies that $\tilde{F}_n(t) \to F(t)$ uniformly a.s. $t \in [0,T)$, as $n \to \infty$.

<u>Case 1c</u>.

$\beta < T$.

Let $\beta \leq x < T$ and $\varepsilon > 0$ be given. By the continuity of F there exists a $0 < z < \beta$ such that $1-F(z) \leq \varepsilon$, and by (5.24) there exists a $N = N(z,\varepsilon)$ such that $F(z) - \varepsilon \leq \tilde{F}_n(z)$, $n \geq N$. Hence, for $n \geq N$, $1 - 2\varepsilon \leq F(z) - \varepsilon \leq \tilde{F}_n(z) \leq \tilde{F}_n(x) \leq 1$. Therefore,

$\lim_{n\to\infty} \tilde{F}_n(x) \to F(x) = 1$ a.s. for $x \geq \beta$. Using (5.24) we have $\tilde{F}_n(x) \to F(x)$ a.s. for $x\varepsilon[0,T)$. Using Polya's theorem again we may conclude that $\tilde{F}_n(x) \to F(x)$ uniformly a.s. for $x\varepsilon[0,T)$, as $n \to \infty$.

Case 2.

F takes a jump on $[0,T)$.

Since F takes a jump on $[0,T)$ at β, it follows that with probability one $K_m = 1$, for some $m = 1,2,\ldots$. Thus, $\tilde{r}_n(t) = \infty$, $\beta \leq t < T$, $n \geq m$, which implies that $\tilde{F}_n(t) = 1$, $\beta \leq t < T$, $n \geq m$. Since $F(t) = 1$, $t \geq \beta$, we have

$$\tilde{F}_n(t) \to F(t) \text{ uniformly a.s. } t \varepsilon [\beta,T), \text{ as } n \to \infty. \tag{5.27}$$

We will now show that the convergence is uniform on $[0,T)$. For $n \geq m$, $\tilde{F}_n$ is a sequence of nondecreasing, bounded, continuous functions. Hence, they may be extended to $[0,\beta]$ in a fashion which will preserve continuity. Similarly, we may extend F to $[0,\beta]$ in a continuous manner. By (5.23), $\tilde{F}_n(t) \to F(t)$, $t\varepsilon[0,\beta)$, as $n\to\infty$. Let $F_n^*(\beta)$, $F^*(\beta)$, be the extended values of F_n and F for $n \geq m$. It is straightforward to show that $F_n^*(\beta) \to F^*(\beta)$ a.s. as $n \to \infty$. Now, let

$$H_n(x) = \begin{cases} \dfrac{F_n(x)}{F_n^*(\beta)} & 0 \leq x < \beta, \\ 1 & x \geq \beta, \\ 0 & x < 0 \end{cases}$$

$$H(x) = \begin{cases} \frac{F(x)}{F^*(\beta)} & 0 \le x < \beta. \\ 1 & x \ge \beta, \\ 0 & x < 0 \end{cases}$$

for $n \ge m$. Note, then, for $n \ge m$, H_n, H are continuous distribution functions, and $H_n(x) \to H(x)$ a.s. Applying Polya's theorem again we may conclude that $H_n(x) \to H(x)$ uniformly a.s., $x \in (-\infty,\infty)$, as $n \to \infty$. Then

$$\tilde{F}_n(x) \to F(x) \text{ uniformly a.s. } x \in [0,\beta), \text{ as } n \to \infty. \tag{5.28}$$

Thus, (5.27) and (5.28) give the desired result. This completes the proof.

We now give two useful corollaries of Theorems 5.10 and 5.11.

Corollary 5.12

Let $S = [u,v]$ be a closed interval of continuity of r, $0 \le u < v < T$. Then, $\tilde{r}_n(x) \to r(x)$ uniformly a.s. on S as $n \to \infty$.

Proof.

By Theorem 5.10

$$\tilde{r}_n(x) \to r(x) \text{ a.s. on } S \text{ as } n \to \infty. \tag{5.29}$$

Case 1.

$$r(u) = r(v) \ge 0$$

For N sufficiently large

$$r(u) - \varepsilon \leq \tilde{r}_n(u) \leq r(u) + \varepsilon$$

and

$$r(v) - \varepsilon \leq \tilde{r}_n(v) \leq r(v) + \varepsilon.$$

But

$$\tilde{r}_n(u) \leq \tilde{r}_n(x) \leq \tilde{r}_n(v).$$

The result follows.

Case 2.

$$r(v) > r(u).$$

Let

$$D_n(x) = \begin{cases} \dfrac{\tilde{r}_n(x) - \tilde{r}_n(u)}{\tilde{r}_n(v) - \tilde{r}_n(u)} & u \leq x \leq v \\ 0 & x < u \\ 1 & x > v \end{cases}$$

$n = 1,2,\ldots,$ and let

$$D(x) = \begin{cases} \dfrac{r(x) - r(u)}{r(v) - r(u)} & u \leq x \leq v \\ 0 & x < u \\ 1 & x > v \end{cases}$$

Note that D_n, D are distribution functions and D is continuous. Applying Polya's theorem (Eisen (1969)) and (5.29) gives

$D_n(x) \to D(x)$ uniformly a.s. on $(-\infty,\infty)$ as $n \to \infty$. For $x \in S$, the sequence $\{r_n(x)\}$, $n \geq 1$ is ultimately uniformly bounded. This implies that $\tilde{r}_n(x) \to r(x)$ uniformly a.s. on S as $n \to \infty$. The proof is completed.

Corollary 5.13

Let $S = [u,v]$ be a closed interval of continuity of r, $0 \leq u < v < T$. Then,

$$\tilde{f}_n(x) \to f(x) \text{ uniformly a.s. on } S \text{ as } n \to \infty$$

where $\tilde{f}_n(x) = \tilde{r}_n(x) \exp(-\int_0^x \tilde{r}_n(y)dy)$.

Proof:

The proof follows directly from Theorem 5.11 and Corollary (5.12).

Next page is blank.

REFERENCES

1. Bray, T. A., Crawford, G. R., and Proschan, F. (1967). Maximum Likelihood Estimation of a U-Shaped Failure Rate Function. Boeing Document D1-82-0660.

2. Brunk, H. D. (1968). On the estimation of parameters restricted by inequalities, Ann. Math. Statist. 29, 437-454.

3. Cramér, Harald (1946). Mathematical Methods of Statistics, Princeton University Press, Princeton.

4. Eisen, Martin (1969). Introduction to Mathematical Probability Theory, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

5. Epstein, Benjamin (1959). Statistical Techniques in Life Testing, Wayne State University.

6. Gnedenko, B. V., Belyayev, Yu K., and Solovyev, A. D. (1969). Mathematical Methods of Reliability Theory, Academic Press, New York.

7. Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters, Ann. Math. Statist. 27, 887-906.

8. Loève, Michel (1963). Probability Theory, D. Van Nostrand Co., Inc., New York.

9. Marshall A. and Proschan, F. (1965). Maximum likelihood estimation for distributions with monotone failure rate, Ann. Math. Statist. 36, 69-77.

10. Rao, Radhakrishna C. (1968). Linear Statistical Inference and Its Applications, John Wiley & Sons, Inc., New York.

11. Wald, A. (1944). On cumulative sume of random variables, Ann. Math. Statist. 15, 283-296.

12. Zelen, M., and Dannemiller, M. C. (1961). The robustness of life testing procedures derived from the exponential distribution, Technometrics 3, 29-49.

Next page is blank.